Rolling A Die Until The First 1

Lazar Ilic

October 2, 2020

Abstract

We show that 6 is the expected number of rolls of a fair die until the first 1 as a consequence of a more general result about a stochastic process.

The problem of determining the expected number of rolls of a fair die until the first 1 is commonly used as a segue to introduce the ideas of Markov chains and state based analysis or arithmetico-geometric series. As it happens, there is another solution.

1 Solving for the Expected Number of Rolls

Possibly the most canonical solution is to set x to be the expected number of rolls and observe the following linear equation recursively:

$$\begin{aligned} x &= \frac{1}{6} + \frac{5}{6}(x+1) \\ \left(\frac{1}{6}\right) x &= 1 \\ x &= 6 \end{aligned}$$

Indeed, the probability of stopping after 1 roll is $\frac{1}{6}$, and, with probability $\frac{5}{6}$, we recurse to the same state after 1 more roll. From this state we again have the same expected number of future rolls.

2 Arithmetico-Geometric Series

By performing casework on the number of rolls, and multiplying by the corresponding probabilities we see that the expected number of rolls is:

$$\begin{aligned}
&\frac{1}{6} + 2\left(\frac{5}{6^2}\right) + 3\left(\frac{5^2}{6^3}\right) + \cdots \\
&\quad = \left(\frac{1}{6} + \frac{5}{6^2} + \cdots\right) + \left(\frac{5}{6^2} + \frac{5^2}{6^3} + \cdots\right) + \left(\frac{5^2}{6^3} + \frac{5^3}{6^4} + \cdots\right) + \cdots \\
&\quad = \frac{\frac{1}{6}}{1-\frac{5}{6}} + \frac{\frac{5}{6^2}}{1-\frac{5}{6}} + \frac{\frac{5^2}{6^3}}{1-\frac{5}{6}} + \cdots \\
&\quad = 6\left(\frac{1}{6} + \frac{5}{6^2} + \frac{5^2}{6^3} + \cdots\right) \\
&\quad = 6\left(\frac{\frac{1}{6}}{1-\frac{5}{6}}\right) = 6
\end{aligned}$$

3 The Process

As in the noted Bernoulli process [1], we have an infinite sequence of binary random variables $X_1, X_2, X_3, \ldots$ which take on the values 0 or 1, except they have associated probabilities $p_1, p_2, p_3, \ldots$ where $X_i = 1$ with probability p_i, and is 0 otherwise. Firstly, we note the expected value of $L = p_1 + p_2 + \cdots + p_n$, where X_n is the first variable that takes on the value 1. If there is a positive probability that none of the $X_i = 1$ then we set $L = p_1 + p_2 + p_3 + \cdots$. This is equivalent with flipping weighted coins, and adding their weights to our sum, until stopping at our first head, and computing the expected value of that sum. Or, humorously, flipping coins forever.

Theorem 1.

$$\begin{aligned}
\mathbb{E}[L] &= p_1 + p_2(1-p_1) + p_3(1-p_2)(1-p_1) + \cdots \\
&= P[\text{stopping}] \leq 1
\end{aligned}$$

Proof. Indeed, the contribution from each p_i to the expected value of this sum is the same as the probability of stopping at the variable X_i because both are the probability we do not stop before the variable multiplied by the probability associated with the variable. □

Now we get the result from before:

Corollary 1. *The expected number of rolls of a fair die until the first* 1 *is* 6.

Proof. Setting $p_i = \frac{1}{6}$ for all i we see that

$$\mathbb{E}[\text{number of rolls}] = 6\mathbb{E}[L] = 6P[\text{stopping}] = 6 \cdot 1 = 6$$

Indeed, $\lim_{n\to\infty}\left(1-\left(\frac{5}{6}\right)^n\right)=1$ and $L=p_1+p_2+\cdots+p_n=\frac{n}{6}$ acts as a counter of the number of rolls so $n=6L$. $\square$

We recover the expected value of the number of trials until the first success in the Bernoulli process similarly. The aforementioned techniques also generalize to this setting. A natural question to ask is: when will this process stop with probability 1?

Theorem 2. *This process stops with probability* 1 *if any of the* p_i *are* 1 *or the infinite sum* $p_1+p_2+p_3+\cdots$ *diverges.*

Proof. That the former leads to stopping is immediate. If the sum diverges and there is positive probability on not stopping then we would not have the inequality above. If the sum converges and none of the probabilities are 1 then there is a positive probability of not stopping. Indeed, set $p_1+p_2+p_3+\cdots=a+b$ where $a\in\mathbb{Z}$ and $b\in[0,1)$ and smooth to the case where $p_i=1-\epsilon$ for $1\le i\le a$ and $p_{a+1}=b+a\epsilon$ for $\epsilon>0$ and note that the resultant probability of not stopping is $\epsilon^a(1-(b+a\epsilon))>0$. We may smooth to this case by shifting the probabilities as needed and noting that $c<d$ implies $cd>(c-e)(d+e)$ where $e>0$ and these are all sensible probabilities. $\square$

Another natural question to ask is: what is the variance of L?

$$\begin{aligned}\mathrm{Var}(L)&=\mathbb{E}[(L-\mathbb{E}[L])^2]\\&=(1-\mathbb{E}[L])[(p_1+p_2+p_3+\cdots)-\mathbb{E}[L]]^2\\&\quad+\sum\nolimits_{i=1}^{\infty}(p_i((1-p_{i-1})(1-p_{i-2})\cdots))[(p_1+\cdots+p_i)-\mathbb{E}[L]]^2\end{aligned}$$

4 Conclusion

Further avenues for research include thinking about structures like this, perhaps even one day the relationship amongst them, equipped with assumptions about the probabilities and going in the direction of known things about credence functions and updating on observations, continuous or discrete in nature. A natural motivation for this process is to ask yourself when you are flipping the coins, when I flip the first head, what is the expected number of heads I ought to have flipped up to and including this flip?

References

[1] Wikipedia, https://en.wikipedia.org/wiki/Bernoulli_process